



SEQUENCE LISTING
SEQUENCE LISTING

<110> Tchaga, Grigory S.
Jokhadze, George

<120> Metal Ion Affinity Tags and Methods for
Using the Same

<130> CLON-056CIP

<140> US 09/858,332

<141> 2001-05-15

<150> 09/404,017

<151> 1999-09-23

<150> 60/101,867

<151> 1998-09-25

<160> 27

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> affinity peptide

<400> 1

His Leu Ile His Asn Val His Lys Glu Glu His Ala His Ala His Asn
1 5 10 15

<210> 2

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> affinity peptide

<400> 2

His Asp Asp His Asp Asp His Asp Asp His Asp Asp His
1 5 10 15

Asp Asp

<210> 3

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> affinity peptide

<400> 3

His Glu Glu His Glu Glu His Glu Glu His Glu Glu His
Page 1

SEQUENCE LISTING
10

1
Glu Glu

5

15

<210> 4
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> affinity peptide

<400> 4
His Asp Glu His Asp Glu His Glu Asn His Glu Asn His Glu Asp His
1 5 10 15
Glu Asp

<210> 5
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> affinity peptide

<400> 5
His Glu Asp His Glu Asp His Glu Asp His Glu Asp His
1 5 10 15
Glu Asp

<210> 6
<211> 5
<212> PRT
<213> Artificial

<220>

<400> 6
Asp Asp Asp Asp Lys
1 5

<210> 7
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> enterokinase cleavage site

<400> 7
Ile Glu Gly Arg
1

<210> 8
<211> 6
<212> PRT

SEQUENCE LISTING

<213> Artificial Sequence

<220>

<223> a factor Xa cleavage site

<400> 8

Leu Val Pro Arg Gly Ser
1 5

<210> 9

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> a thrombin cleavage site

<400> 9

His Pro Phe His Leu Val Ile His
1 5

<210> 10

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> a renin cleavage site

<400> 10

Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
1 5 10

<210> 11

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> an immunological tag

<400> 11

Asp Tyr Lys Asp Asp Asp Asp Lys
1 5

<210> 12

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> an immunological tag

<400> 12

Cys Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
1 5 10

<210> 13

SEQUENCE LISTING

<211> 3426

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA sequence of vector containing cDNA of recombinant enterokinase

<400> 13

gacgaaagg	cctcgtgata	cgccctat	tttt	tataggttaa	tgatcatgata	ataatggttt	60
cttagacgtc	aggtggcact	tttcggggaa	atgtg	atgtg	atgtg	atgtg	120
tctaaataca	ttcaaatatg	tatccgctca	tgagacaata	accctgataa	atgcttcaat	180	
aatattgaaa	aaggaagagt	atgagtattc	aacatttccg	tgatcgccctt	attccctttt	240	
ttgcggcatt	ttgccttcct	gtttttgctc	acccagaaac	gctgggtgaaa	gtaaaagatg	300	
ctgaagatca	gttgggtgca	cgagtgggtt	acatcgaaact	ggatctcaac	agcggtaaga	360	
tccttgagag	ttttcgcccc	gaagaacgtt	ttccaatgat	gagcactttt	aaagtctctg	420	
tatgtggcgc	ggtattatcc	cgtattgacg	ccgggcaaga	gcaactcggg	cgccgcatac	480	
actatttctca	gaatgacttg	gttgagtact	caccagtcac	agaaaagcat	cttacggatg	540	
gcatgacagt	aagagaatta	tgacgtgctg	ccataaccat	gagtataaac	actgcggcca	600	
acttacttct	agttatctac	ggaggaccga	aggagctaac	cgcttttttg	cacaacatgg	660	
gggatcatgt	aactcgcctt	gatcggtggg	agcattggga	gaatgaagcc	ataccaaacg	720	
acgagcgtga	caccacgatg	cctgtagcaa	tggcaacaac	gttgcgcaaa	ctattaactg	780	
gcgaactact	tactctagct	tcccggcaac	aattaataga	ctggatggag	gcggataaag	840	
ttgcaggacc	acttctgcgc	tcggcccttc	cggctggctg	gtttattgct	gataaatctg	900	
gagccgggtga	gcgtgggtct	cgcggtatca	ttgcagcact	ggggccagat	ggtaagccct	960	
cccgtatcgt	agttatctac	acgacgggga	gtcaggcaac	tatggatgaa	cgaaaatagac	1020	
agatcgctgt	gataggtgcc	tcactgatta	agcattggga	actgtcagac	caagtttact	1080	
catatatact	ttagattgat	ttaaaacttc	atttttaatt	taaaaggatc	taggtgaaga	1140	
tcctttttga	taatctcatg	acccaaatcc	cttaacgtga	gttttcgttc	cactgagggt	1200	
cagaccccg	agaaaagatc	aaaggatcct	cttgagatcc	tttttttctg	cgcgtaattct	1260	
gctgcttgca	aacaaaaaaa	ccaccgctac	cagcgggtgg	ttgtttgccg	gatcaagagc	1320	
taccaactct	ttttccgaag	gtaactggct	tcagcagagc	gcagatacca	aatactgtcc	1380	
ttctagtgtg	gccgtagtta	ggccaccact	tcaagaactc	tgtagcaccg	cctacatacc	1440	
tcgctctgct	aatcctgtta	ccagtggctg	ctgccagtgg	cgataagtcg	tgtcttaccg	1500	
ggttggactc	aagacgatag	ttaccggata	aggcgcagcg	gtcgggctga	acgggggggt	1560	
cgtgcacaca	gcccagcttg	gagcgaacga	cctacaccga	actgagatac	ctacagcgtg	1620	
agctatgaga	aagcgccacg	cttcccgaag	ggagaaaagg	ggacagggtat	ccggtaagcg	1680	
gcagggtcgg	aacaggagag	cgacagagg	agcttccagg	gggaaacgcc	tggtatcttt	1740	
atagtcctgt	cggttttcgc	cacctctgac	ttgagcgtcg	atttttgtga	tgctcgtcag	1800	
gggggaggag	cctatggaaa	aacgcccagc	acgcccctt	tttacgggtc	ctggcctttt	1860	
gctggccttt	tgctcacatg	ttctttcctg	cgttatcccc	tgattctgtg	gataaccgta	1920	
ttaccgcctt	tgagttagct	gataccgctc	gccgcagccg	aacgaccgag	cgagcagagt	1980	
cagttagcga	ggaagcggaa	gagcgcccaa	tacgcaaacc	gcctctcccc	gcgcgttggc	2040	
cgattcatta	atgcagctgg	cacgacaggt	ttcccagact	gaaagcgggc	agttagcgca	2100	
acgcaattaa	tgtgagttag	ctcactcatt	aggcacccca	ggctttacac	tttatgcttc	2160	
cggctcgtat	gttggtgtga	attgtgagcg	gataacaatt	tcacacagga	aacagctatg	2220	
accatgatta	cgccaagcct	gaaggatcat	ctcatccaca	atgtccacaa	agaggagcac	2280	
gctcatgccc	acaacaagat	cgatattgtc	ggaggaaagt	actccagaga	aggagcctgg	2340	
ccttgggtcg	ttgctctgta	tttcgacgat	caacaggctc	gcggagcttc	tctggtgagc	2400	
agggattggc	tggtgtcgag	cgcccactgc	gtgtacggga	gaaatatgga	gccgtctaag	2460	
tggaaagcag	tgctaggcct	gcataatggc	tcaaatctga	cttctcctca	gatagaaact	2520	
aggttgattg	accaaattgt	cataaaacca	cactacaata	aacggagaaa	gaacaatgac	2580	
attgccatga	tgcattctga	aatgaaagtg	aactacacag	attatatata	gcctatttgt	2640	
ttaccagaag	aaaatcaagt	ttttccccc	ggaagaattt	gttctattgc	tggctggggg	2700	
gcacttatat	atcaaggttc	tactgcagac	gtactgcaag	aagctgacgt	ttcccttcta	2760	
tcaaatgaga	aatgtcaaca	acagatgcc	gaatataaca	ttacggaaaa	tatggtgtgt	2820	
gcaggctatg	aagcaggagg	ggtagattct	tgtagggggg	attcaggcgg	accactcatg	2880	
tgccaagaaa	acaacagatg	gctcctggct	ggcgtgacgt	catttggata	tcaatgtgca	2940	
ctgcctaata	gcccaggggt	gtatgcccgg	gtcccaagg	tcacagagtg	gatacaaagt	3000	
tttctacatg	agctcgtaat	tagctgagaa	ttcactggcc	gtcgttttac	aacgtcgtga	3060	
ctgggaaaaa	cctggcgtaa	cccaacttaa	tcgctttgca	gcacatcccc	ctttcgccag	3120	
ctggcgtaat	acggaagagg	ccgcacagg	tcgccccttc	caacagttgc	gcagcctgaa	3180	
tggcgaatgg	cgccctgatgc	ggtattttct	ccttacgcat	ctgtgcggta	tttcacaccg	3240	

SEQUENCE LISTING

catatggtgc actctcagta caatctgctc tgatgccgca tagttaagcc agcccccgaca 3300
 cccgcccaaca cccgctgacg cgccctgacg ggcttgctctg ctcccggcat ccgcttacag 3360
 acaagctgtg accgtctccg ggagctgcat gtgtcagagg ttttcaccgt catcaccgaa 3420
 acgcgc 3426

<210> 14
 <211> 269
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> protein sequence of vector containing cDNA of recombinant enterokinase

<400> 14
 Met Thr Met Ile Thr Pro Ser Leu Lys Asp His Leu Ile His Asn Val
 1 5 10 15
 His Lys Glu Glu His Ala His Ala His Asn Lys Ile Asp Ile Val Gly
 20 25 30
 Gly Ser Asp Ser Arg Glu Gly Ala Trp Pro Trp Val Val Ala Leu Tyr
 35 40 45
 Phe Asp Asp Gln Gln Val Cys Gly Ala Ser Leu Val Ser Arg Asp Trp
 50 55 60
 Leu Val Ser Ala Ala His Cys Val Tyr Gly Arg Asn Met Glu Pro Ser
 65 70 75 80
 Lys Trp Lys Ala Val Leu Gly Leu His Met Ala Ser Asn Leu Thr Ser
 85 90 95
 Pro Gln Ile Glu Thr Arg Leu Ile Asp Gln Ile Val Ile Asn Pro His
 100 105 110
 Tyr Asn Lys Arg Arg Lys Asn Asn Asp Ile Ala Met Met His Leu Glu
 115 120 125
 Met Lys Val Asn Tyr Thr Asp Tyr Ile Gln Pro Ile Cys Leu Pro Glu
 130 135 140
 Glu Asn Gln Val Phe Pro Pro Gly Arg Ile Cys Ser Ile Ala Gly Trp
 145 150 155 160
 Gly Ala Leu Ile Tyr Gln Gly Ser Thr Ala Asp Val Leu Gln Glu Ala
 165 170 175
 Asp Val Pro Leu Leu Ser Asn Glu Lys Cys Gln Gln Gln Met Pro Glu
 180 185 190
 Tyr Asn Ile Thr Glu Asn Met Val Cys Ala Gly Tyr Glu Ala Gly Gly
 195 200 205
 Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Met Cys Gln Glu
 210 215 220
 Asn Asn Arg Trp Leu Leu Ala Gly Val Thr Ser Phe Gly Tyr Gln Cys
 225 230 235 240
 Ala Leu Pro Asn Arg Pro Gly Val Tyr Ala Arg Val Pro Arg Phe Thr
 245 250 255
 Glu Trp Ile Gln Ser Phe Leu His Glu Leu Val Ile Ser
 260 265

<210> 15
 <211> 12
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> an amino acid sequence embodiment of the affinity purification site

<400> 15
 His Asn His Asn His Asn His Asn His Asn His Asn
 1 5 10

SEQUENCE LISTING

<210> 16
 <211> 48
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> a DNA sequence embodiment of the affinity
 purification site

 <400> 16
 catctcatcc acaatgtcca caaagaggag cacgctcatg cccacaac 48

 <210> 17
 <211> 36
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> a DNA sequence embodiment of the affinity
 purification site

 <400> 17
 cataaccata accataacca taaccataac cataac 36

 <210> 18
 <211> 54
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> a DNA sequence embodiment of the affinity
 purification site

 <400> 18
 catgatgatc atgatgatca tgatgatcat gatgatcatg atgatcatga tgat 54

 <210> 19
 <211> 54
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> a DNA sequence embodiment of the affinity
 purification site

 <400> 19
 catgaggagc atgaggagca tgaggagcat gaggagcatg aggagcatga ggag 54

 <210> 20
 <211> 54
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> a DNA sequence embodiment of the affinity
 purification site

 <400> 20
 catgatgagc atgatgagca tgagaacat gagaacatg aggatcatga ggat 54

 <210> 21

SEQUENCE LISTING

<211> 9
<212> PRT
<213> Human

<220>
<221> VARIANT
<222> 9
<223> Xaa at position 9 is an amino acid with an
aliphatic or amide side chain.

<221> VARIANT
<222> 2
<223> Xaa at position 2 is an amino acid with an
aliphatic or amide side chain

<221> VARIANT
<222> 3
<223> Xaa at position 3 is an amino acid with an
aliphatic or amide side chain.

<221> VARIANT
<222> 5
<223> Xaa at position 5 is an amino acid with a basic
side chain (except HIS) or an acidic side chain.

<221> VARIANT
<222> 6
<223> Xaa at position 6 is an amino acid with a basic
side chain (except HIS) or an acidic side chain.

<221> VARIANT
<222> 7
<223> Xaa at position 7 is an amino acid with a basic
side chain (except HIS) or an acidic side chain.

<400> 21
His Xaa Xaa His Xaa Xaa His Xaa
1 5

<210> 22
<211> 6
<212> PRT
<213> Human

<220>

<400> 22
His Arg His Arg His Arg
1 5

<210> 23
<211> 9
<212> PRT
<213> Human

<220>
<221> VARIANT
<222> 2
<223> Xaa = an amino acid having an acidic side chain.

<221> VARIANT

SEQUENCE LISTING

<222> 3
<223> Xaa = an amino acid having an acidic side chain.

<221> VARIANT
<222> 5
<223> Xaa = an amino acid having an acidic side chain.

<221> VARIANT
<222> 6
<223> Xaa = an amino acid having an acidic side chain.

<221> VARIANT
<222> 8
<223> Xaa = an amino acid having an acidic side chain.

<221> VARIANT
<222> 9
<223> Xaa = an amino acid having an acidic side chain.

<400> 23
His Xaa Xaa His Xaa Xaa His Xaa Xaa
1 5

<210> 24
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Protease cleavage site

<400> 24
Asp Asp Asp Asp Lys
1 5

<210> 25
<211> 21
<212> PRT
<213> Artificial Sequence

<220>
<223> Translation of nucleotide coding sequence

<400> 25
Leu Pro Pro Leu Ser Glu Leu Ile Pro Leu Ala Ala Ala Glu Arg Pro
1 5 10 15
Ser Ala Ala Ser Gln
20

<210> 26
<211> 38
<212> PRT
<213> Artificial Sequence

<220>
<223> Translation of nucleotide coding sequence

<400> 26
Ala Arg Lys Arg Lys Ser Ala Gln Tyr Ala Asn Arg Leu Ser Pro Arg
1 5 10 15

SEQUENCE LISTING

Val Gly Arg Phe Ile Asn Ala Ala Gly Thr Thr Gly Phe Pro Thr Gly
 20 25 30
 Lys Arg Ala Val Ser Ala
 35

<210> 27

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Translation of nucleotide coding sequence

<400> 27

Glu Phe Thr Gly Arg Arg Phe Thr Thr Ser
 1 5 10